ANL-7667

ANL-7667

Argonne National Laboratory

C:REGUSE, A Linear and Quadratic Regression User Program

by

Conrad E. Thalmayer

Printed in the United States of America
Available from
Clearinghouse for Federal Scientific and Technical Information
National Bureau of Standards, U. S. Department of Commerce
Springfield, Virginia 22151
Price: Printed Copy $3.00; Microfiche $0.65

ANL-7667
Mathematics and
Computers

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439

C:REGUSE,
A Linear and Quadratic Regression User Program

by

Conrad E. Thalmayer

Chemistry Division

February 1970

PREFACE

C:REGUSE is written in XDS SYMBOL to run as a "Background" program under ARGOS (Argonne Operating System), using the C:READ, C:WRITE, and C:REGN routines. The limitation upon the number of data points is due only to the size of the Background core area; when more Background area is available, the program can be readily changed to treat more data.

This report presents all the information needed for the use and understanding of C:REGUSE; the derivation of the regression equations, details of use of the called routines, and their organization are given in various ANL Reports.[1-3] Section I of this report explains the purpose of this program and the significance of its results. Section II contains the instructions for its use. Section III, together with the flow charts and listing, indicates the structure of the program and will facilitate any modification by the user.

C:REGUSE was written in December 1968.

TABLE OF CONTENTS

	Page
ABSTRACT	5
I. LEAST SQUARES, GENERAL	5
II. USE OF THE PROGRAM	6
III. ORGANIZATION	7
APPENDIX A. Flow Charts	12
APPENDIX B. Listing	14
ACKNOWLEDGMENTS	19
REFERENCES	19

LIST OF FIGURES

No.	Title	Page
1.	Sample Input with Three Sets of Data.	8
2.	Output for the Data of Fig. 1	9
3.	Data of Fig. 1 and Curves Calculated from the Parameters of Fig. 2.	11

C:REGUSE,
A Linear and Quadratic Regression User Program

by

Conrad E. Thalmayer

ABSTRACT

This report describes a user program for the Xerox Data Systems Sigma 5 or Sigma 7 computer operating under ARGOS (Argonne Operating System). C:REGUSE performs a linear or quadratic weighted least-squares analysis upon not more than 500 data points, yielding the coefficients of the fitted equation, their internal errors, the external errors, and the chi-square value; it uses card input and line-printer output. The report explains the significance of the program, describes its use, and outlines its structure. The flow charts and listing are included.

I. LEAST SQUARES, GENERAL*

Probably the most common statistical problem in science is presented by a series of measurements of two associated variables (x,y). It is assumed that the observed (x,y) association results from an underlying relationship and that this relationship can be expressed in the form of a curve. Observed deviations from this curve are then considered as statistical fluctuations, possibly resulting from experimental errors in measurement or from inherent fluctuations. The problem is to estimate a "true" curve which gives the best possible fit to the data, despite the statistical fluctuations.

The method of least squares is particularly suitable for this task, for (a) it is objective; (b) it provides definite weights; (c) it permits a test for checking the suitability of the functional form chosen; (d) it yields the errors in the curve estimates based on the stated errors in the measurements (internal errors) and also the errors resulting from the imperfect fit of the data to the curve (external errors); (e) it is applicable to cases with more than two variables. Most importantly, (f) least squares gives high-quality estimates, i.e., the results lie closer to the "true" values, on the average, than do the results of graphical fitting or most other methods of calculation.

---

*This treatment is adapted from the work of Jaffey.[4]

C[illegible]
A Linear and Quadratic Regression
User Program [illegible]

by

[illegible]

ABSTRACT

This report describes a [illegible] for the Xerox Data Systems Sigma 7 computer operating under ARCOS (Argonne [illegible] System). [illegible] performs a linear or quadratic weighted least squares analysis [illegible] not more than 500 [illegible] points, yielding the coefficients of the fitted equation, their [illegible] errors, the external errors, and the chi-square value [illegible] input and line printer output. The report explains the significance of the program, describes its use, and [illegible] structure. The flow charts and listing are included.

I. LEAST SQUARES IN GENERAL*

Probably the most common statistical problem in [illegible] by a series of measurements of two correlated variables. It is assumed that the observed correlation results from an underlying [illegible] whether this relationship can be expressed in the form of a curve. Observed deviations from that curve are then considered as statistical fluctuations possibly resulting from experimental errors or inadequacies from inherent [illegible]. The problem is to estimate a "true" curve which gives the best possible fit to the data despite the statistical fluctuations.

The method of least squares is particularly suitable for this task [illegible] (b) it provides definite weighting [illegible] for checking the suitability of the functional form chosen, (d) it yields the errors [illegible] of the stated errors [illegible] external errors [illegible] the data. [illegible] it is applicable to cases [illegible] the "large" values [illegible] other methods [illegible]

*[illegible] adapted from the work of [illegible]

There is no automatic way to select the appropriate analytic form for the curve to be fitted. Experience in the science, graphical analysis, and theoretical guidance seem to be the best sources for suggesting a suitable type of curve. Simplicity is a necessary criterion, for n measurements may always be exactly fitted with an n-parameter function. The number of parameters is chosen to be the minimum required to reduce deviations of the measurements from the curve to an acceptable minimum. Wise choice of functional form is essential, e.g., data adequately fitted with a two-parameter exponential curve may require many more parameters if a polynomial is fitted. The method of least squares evaluates parameters after the curve type has been selected on nonstatistical grounds.

The user of this program assumes that his data, within the limits of his stated standard deviations, will be adequately fitted by the chosen functional form. If the given standard deviations are accurate absolute values, incorporating the errors from all sources, he may test his assumption by the "χ^2 Test." The χ^2 value calculated by this program is compared with the value given in a χ^2 table at a chosen rejection level, say 5%, and (N-M) degrees of freedom, where N is the number of observations and M is the number of parameters in the fitted equation. If the calculated value exceeds the tabulated value, then the chosen functional form is not consistent with the given observations and standard deviations, and should be rejected. (If accurate absolute standard deviations are not available, but replicate y-observations have been made at a representative number of x-values, consistency may be tested by the "F Test," independently of this routine.)

II. USE OF THE PROGRAM

C:REGUSE performs a regression analysis upon not more than 500 sets of values of x_i, y_i, and σ_i, yielding a weighted least-squares fit to either the linear equation $y = a + bx$ or the quadratic equation $y = a + bx + cx^2$, or both in turn. The x_i are assumed to be without error. The error of each y_i is expressed as its standard deviation σ_i. Each y_i is weighted by $w_i = 1/\sigma_i^2$; thus a full array of nonzero σ_i values is required, though their significance may be only relative to each other.

In addition to the equation parameters a and b (and c for the quadratic case), C:REGUSE calculates $\sigma^2(a)$, $\sigma^2(b)$, $s^2(a)$, $s^2(b)$, $s^2(y)$, and χ^2 (and $\sigma^2(c)$ and $s^2(c)$ for the quadratic case). The σ^2 are the internal errors (variances) of the parameters, calculated purely by propagation of the given standard deviations σ_i. (If the standard deviations are only relative, the internal errors are meaningless.) The s^2 are the external errors (estimated variances), which indicate the closeness of fit of the x_i, y_i to the calculated curve. The quantity χ^2 is a measure of the agreement between the external and internal errors; it can be used with a χ^2 table to

There is no automatic way to select the appropriate analytic form for the data to be fitted. Experience in the subject, graphical analysis and theoretical knowledge seem to be the best sources for suggesting a suitable type of curve. Simplicity is a useful criterion for it [illegible] ...the [illegible] fit with [illegible] parameter [illegible]. If the number of parameters is chosen to be too [illegible], the [illegible] deviations of the measurements [illegible] to some [illegible]. Wise choice of functional form is [illegible], e.g., [illegible] within [illegible] parameter exponential [illegible] may require many more parameters [illegible] polynomial [illegible]. The method of least squares estimates parameters after the curve form has been selected on [illegible] grounds.

The user of this program assumes that his data [illegible] of his [illegible] standard deviations will be [illegible] by the chosen functional form. If the given standard deviations [illegible] values, including the errors from all sources, the [illegible] may [illegible] the "χ^2 Test". The χ^2 value calculated [illegible] this program [illegible] compared with the value given in a χ^2 table if a chosen [illegible] level, say 5%, and (N - [illegible]) degrees of freedom, where N is the number of observations and M is the number of parameters in the fitted equation. If the calculated value exceeds the tabulated value, then the chosen functional form is not consistent with the given observations and standard deviations, and should be rejected. If accurate [illegible] standard deviations are not available, but replicate y-observations have been made at a representative number of x-values, consistency may be tested by the "F-Test," independently of this routine [illegible].

[illegible] OF THE PROGRAM

REGOSE performs a regression analysis upon not more than [illegible] sets of values of x_i [illegible], yielding a weighted least-squares fit to [illegible] the linear [illegible] the [illegible]. The x_i are assumed [illegible]. The [illegible] of each y_i [illegible] as its standard deviation [illegible], weighted by [illegible] full array [illegible] required, [illegible] significance may be only [illegible].

In addition to the [illegible] parameters [illegible] REGOSE [illegible] the [illegible] of the parameters [illegible] deviations [illegible] standard [illegible] errors [illegible], which indicate [illegible] calculated curve. The quantity [illegible] measure of [illegible] the external and internal errors [illegible] table.

judge whether the chosen equation form is consistent with the data. The significance of these quantities is explained more fully in Sect. I.

Input format:

(1) A card bearing L, Q, or B in the first column, to indicate linear or quadratic regression, or both. The remaining 79 columns may contain the user's title.

(2) Not more than 500 cards, each containing a set of values of x_i, y_i, and σ_i in free format (separated by spaces).

The above may be repeated any number of times.

(3) A card bearing # in the first column, to indicate the end.

Figure 1 shows the input for a sample run with three sets of data, utilizing the three options. A variety of acceptable number formats is also illustrated. Figure 2 is the output from this run. Figure 3 is a plot of the data with their standard deviations and the fitted curves. (The linear and quadratic curves for Example 3 are so close together that they cannot be distinguished on this scale.)

For Example 1, the calculated χ^2 value of 2.7 is less than the value 6.0 given in a χ^2 table for a 5% rejection level and (N-M) = 2. According to this criterion, therefore, a linear equation is suitable for the data of Example 1. For Example 2, the calculated χ^2 value of 11.4 is greater than 7.8, the tabulated χ^2 value for a 5% rejection level and (N-M) = 3. A quadratic equation is therefore not consistent with this data at a 95% confidence level. For Example 3, the calculated χ^2 value of 5.3 for a linear fit is less than 9.5, the tabulated χ^2 value for 5% rejection and (N-M) = 4. The χ^2 value calculated for a quadratic fit also meets the χ^2 test (5.3 versus 7.8 for 5% and (N-M) = 3). However, the criterion of simplicity now dictates that a linear euqation is more appropriate for these data.

Note added in proof: C:REGUSE has been extended to accept 512 data points and to yield a table of x_i, y_i(given), y_i(calc), σ_i, residual$_i$ = y_i(given) - y_i(calc), and residual$_i$/y_i(calc).

III. ORGANIZATION

As can be seen from the first flow chart and the listing, the program flow depends first on the content of column 1 of each data card. If the C:READ routine interprets the content of column 1 as a valid number character (digit, blank, +, -, ., E, or D), the point count is incremented by 1, the values of x_i, y_i, and σ_i are read by means of C:READ and stored as long floating-point numbers, and the next card is fetched. If the point count is greater than 500 or the fields read for x_i, y_i, and σ_i do not contain valid numbers, the program stops.

as whether the [illegible] equation is [illegible] consistent with the data. The [illegible] difference [illegible] equations is explained [illegible] in Sect. I.

[illegible]

[illegible] and [illegible] Q [illegible] in the first column [illegible] linear, quadratic expression, or [illegible]. The remaining [illegible] columns may contain the [illegible].

(5) [illegible] than 512 cards [illegible] containing a set of values x_i, y_i and σ_i in [illegible] format (or [illegible] spaces).

The above may be repeated [illegible] number of times.

[illegible] card reading [illegible] indicate the [illegible].

Figure 1 shows the input for [illegible] example [illegible] with [illegible] data, utilizing [illegible] options. A variety [illegible] is also illustrated. [illegible] is the output [illegible] plot of the data with their standard deviation [illegible] linear and quadratic curves for Example [illegible] that they cannot be distinguished on this scale.)

[illegible] Example [illegible] the calculated χ^2 value [illegible] less than the value 6.0 [illegible] χ^2 table [illegible] rejection level and [illegible]. According to this [illegible], therefore, a linear equation [illegible] for the data of Example 1. For Example 2, the calculated χ^2 value [illegible] is greater than [illegible] calculated χ^2 value for [illegible] rejection level and (N-M) [illegible]. A quadratic equation is [illegible] consistent with this data at a 95% confidence level. For Example [illegible] calculated χ^2 value of 5.5 for a linear, it is less than 9 [illegible] calculated χ^2 value for [illegible] rejection and (N-M) [illegible]. The χ^2 value calculated for a quadratic also meets the χ^2 test [illegible] versus 7.8 for [illegible] and (N-M) = 3). [illegible] criterion of simplicity now dictates that a linear [illegible] is more appropriate for these data.

[illegible] added [illegible] GREGUSI [illegible] to accept 512 data points and to yield [illegible] (plotted, and residuals [illegible]).

III. OR[illegible]

As can be [illegible] from the first [illegible] and the [illegible] are [illegible] in the contents of [illegible] in [illegible] the [illegible] point [illegible] values of [illegible] and [illegible] read [illegible] numbered and the [illegible] used [illegible] the [illegible] here, the [illegible].

INPUT DATA

```
         1         2         3         4         5         6         7         8
12345678901234567890123456789012345678901234567890123456789012345678901234567890
L JOHN Q. PUBLIC             REGRESSION ANALYSIS              EXAMPLE NUMBER ONE
                     1           9        1
                     3          12        1
                     5          16        1
                     9          19        1
Q JOHN Q. PUBLIC             REGRESSION ANALYSIS              EXAMPLE NUMBER TWO
                     1           5       .2
                     3           4       .2
                     5           5       .2
                     7           6       .2
                     9           9       .3
                    13          15       .4
B JOHN Q. PUBLIC            REGRESSION ANALYSIS             EXAMPLE NUMBER THREE
            1.000       1..          .5
            3.0000      1..0         .5
            50E-1       30E-1        .5
            70.E-1      .3E1         .5
            .9D1        500D-2       .5
            .13E+2      .06E 2       .5
#
12345678901234567890123456789012345678901234567890123456789012345678901234567890
         1         2         3         4         5         6         7         8
```

AMD-8 (9-60)

Fig. 1. Sample Input with Three Sets of Data

```
JØHN Q. PUBLIC          REGRESSIØN ANALYSIS          EXAMPLE NUMBER ØNE          10:55:04          03-31-70



                              LINEAR REGRESSIØN
                                   Y=A+BX

        PARAMETERS                      INTERNAL ERRØRS                      EXTERNAL ERRØRS

 A = 8.3428571428571447 4E00     SIGSQA = 8.2857142857142862E-01     SSQA = 1.1126530612244920 3E00
 B = 1.2571428571428588 9E00     SIGSQB = 2.8571428571428576E-02     SSQB = 3.8367346938775537E-02
                                                                     SSQY = 1.3428571428571454 1E00

                                 CHISQ = 2.6857142857142892 7E00



JØHN Q. PUBLIC          REGRESSIØN ANALYSIS          EXAMPLE NUMBER TWØ          14:00:41          02-23-70

                              QUADRATIC REGRESSIØN
                                  Y=A+BX+CX**2

        PARAMETERS                      INTERNAL ERRØRS                      EXTERNAL ERRØRS

 A = 5.3165085388994310 9E00     SIGSQA = 6.1919829222011431E-02     SSQA = 2.3490283093711294E-01
 B = -6.3544170356314 59E-01     SIGSQB = 8.4140839131351659E-03     SSQB = 3.1920180591119282E-02
 C = 1.0826481130086453E-01      SIGSQC = 5.2445709466582435E-05     SSQC = 1.9896123388896762E-04
                                                                     SSQY = 3.7936608335090349 3E00

                                 CHISQ = 1.1380982500527119 2E01
```

Fig. 2. Output for the Data of Fig. 1

```
JØHN Q. PUBLIC          REGRESSIØN ANALYSIS          EXAMPLE NUMBER THREE                14:00:43        02-23-70



                                          LINEAR REGRESSIØN
                                               Y=A+BX


         PARAMETERS                          INTERNAL ERRØRS                              EXTERNAL ERRØRS

  A = 2.7142857142857101E-01        SIGSQA = 1.4910714285714301E-01         SSQA = 1.9809948979591847E-01
  B = 4.5714285714285729E-01        SIGSQB = 2.6785714285714297E-03         SSQB = 3.5586734693877553E-03
                                                                            SSQY = 1.3285714285714298E00

                                     CHISQ = 5.3142857142857147E00



                                        QUADRATIC REGRESSIØN
                                             Y=A+BX+CX**2


         PARAMETERS                          INTERNAL ERRØRS                              EXTERNAL ERRØRS

  A = 1.7321428571428205E-01        SIGSQA = 3.5167410714285756E-01         SSQA = 6.1738343253968295E-01
  B = 4.9880952380952452E-01        SIGSQB = 3.9136904761904749E-02         SSQB = 6.8707010582010632E-02
  C = -2.9761904761904810E-03       SIGSQC = 1.8601190476190470E-04         SSQC = 3.2655423280423296E-04
                                                                            SSQY = 1.7555555555555566E00

                                     CHISQ = 5.2666666666666668E00
```

Fig. 2 (Contd.)

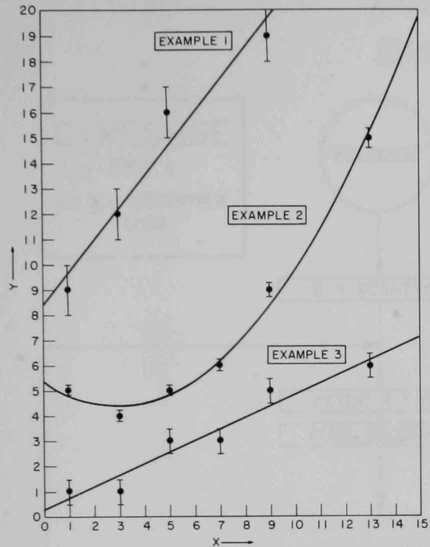


Fig. 3. Data of Fig. 1 and Curves Calculated from the Parameters of Fig. 2

If the content of column 1 is not a valid number character, it is stored in location TYPEBUFF. If a prior batch of data has not been treated (as shown by REGNTYPE still containing its initial value of zero), the program goes to the address TYPESETR. If a prior batch of data has been treated, the program goes to REGNHEAD.

At TYPESETR, the content of TYPEBUFF is loaded into location REGNTYPE. If this content is '#', the program exits normally. Otherwise, the user's title is loaded into an output buffer, the time and date are fetched and loaded into the same buffer, and that line is printed; the point count is now set to zero and the next card is fetched.

At REGNHEAD, as can be seen from the second flow chart, the appropriate heads are printed, depending on whether REGNTYPE contains 'Q'. The appropriate regression analysis is then performed by the C:REGN routine, as determined by the same criterion. The results of the analysis are then converted to EBCDIC decimal by the C:WRITE routine, loaded into their respective output buffers, and printed. Finally, if REGNTYPE contains 'B', this is replaced by 'Q' and the program returns to address REGNHEAD; otherwise, it goes to TYPESETR to treat the next batch of data or exit.

If the content of column 1 is not a valid number character, it is stored in location TYPEBUFF. If a prior batch of data has not been treated (as shown by RECNTYPE still containing the initial value of zero), the program goes to the address TYPESEEK. If a prior batch of data has been treated, the program goes to RECNHEAD.

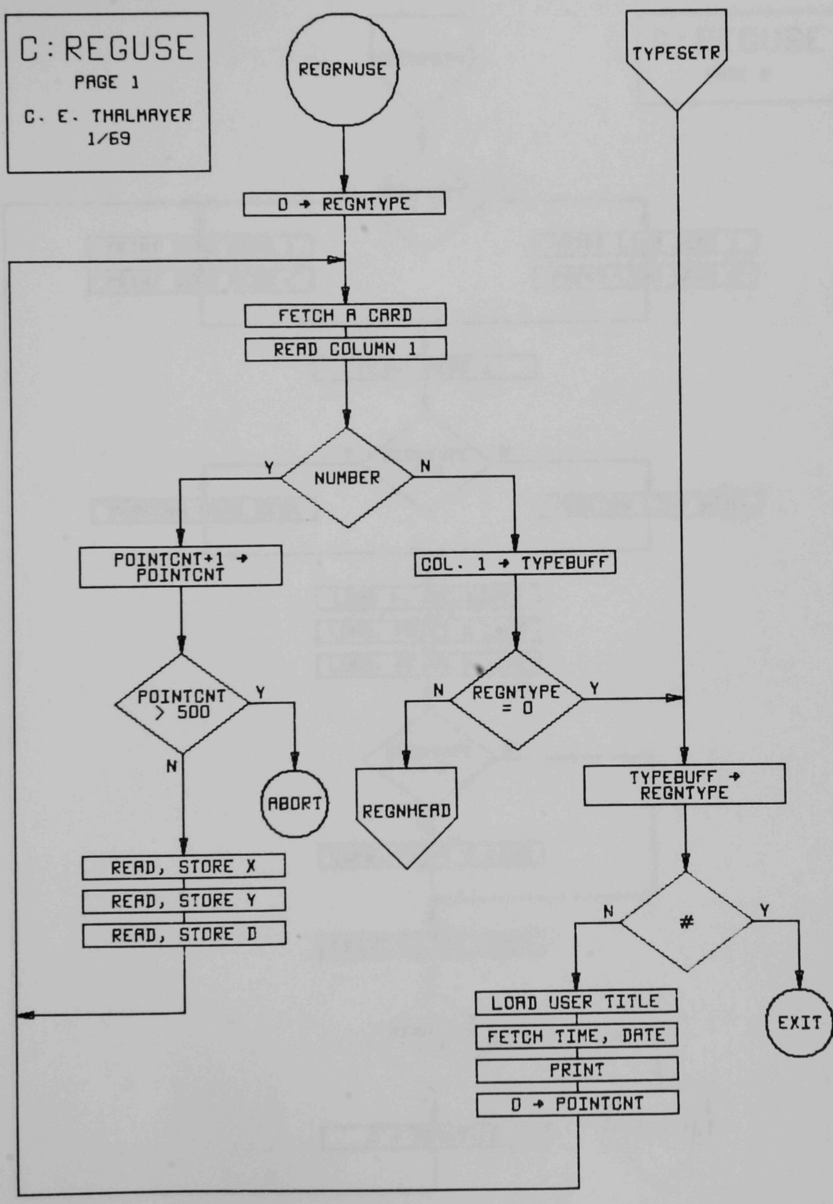
At TYPESEEK, the content of TYPEBUFF is loaded into location RECNTYPE. If this content is [illegible], the program exits normally. Otherwise, [illegible] title is loaded into an output buffer, the time and date are fetched and loaded into the same buffer, and that line is printed. The point count is now set to zero and the next card is fetched.

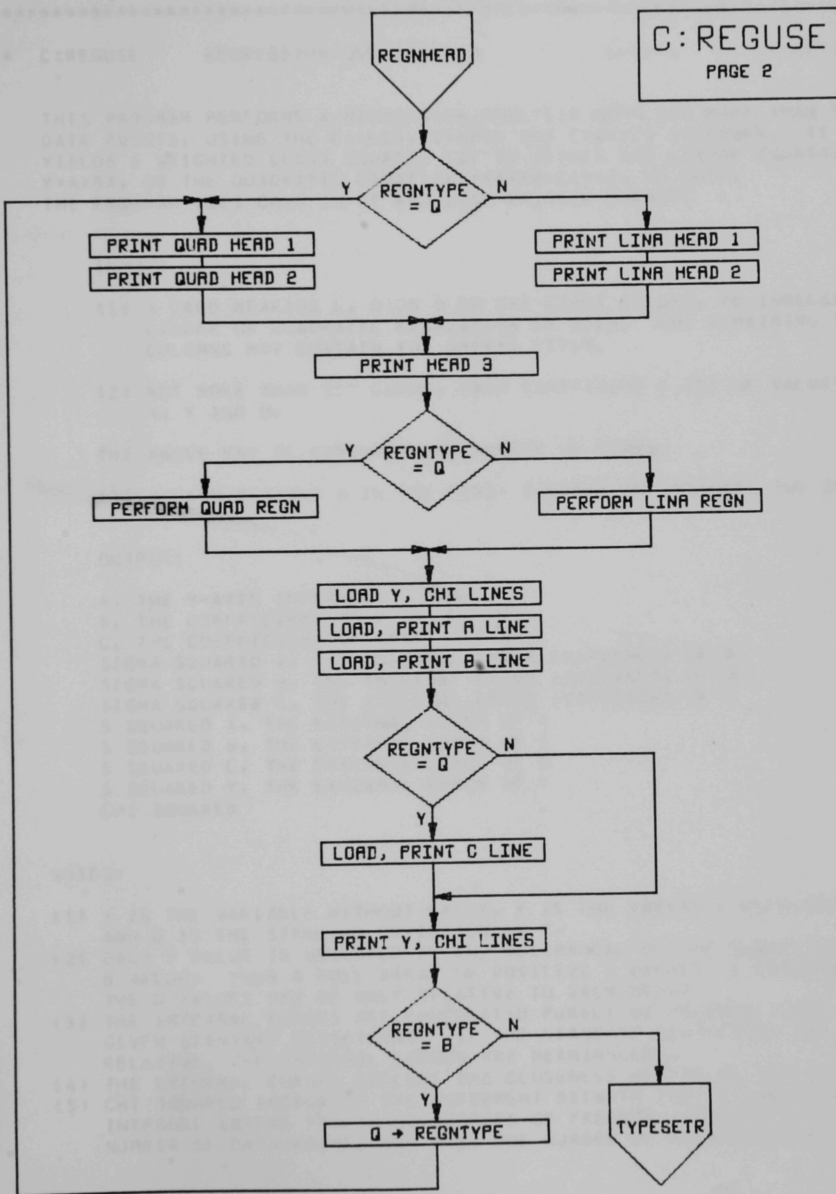
At RECNHEAD, as can be seen from the second flow chart, the appropriate heads are printed depending on whether RECNTYPE contains [illegible]. The appropriate [illegible] analyses is then performed by the [illegible] RECN routine, as determined by the same selection. The results of the analyses are then converted to EBCDIC decimal by the CONVERT routine, loaded into their respective output buffers and printed. Finally, if RECNTYPE contains [illegible], this is replaced by [illegible] and the program returns to address RECNHEAD; otherwise, it goes to TYPESEEK to treat the next batch of data or exit.

Fig. 8. Data of Fig. 1 and Curves Calculated from the Parameters of Fig. 9

APPENDIX A

Flow Charts




C:REGUSE
PAGE 2
REGNHEAD
REGNTYPE = Q
Y
N
PRINT QUAD HEAD 1
PRINT QUAD HEAD 2
PRINT LINA HEAD 1
PRINT LINA HEAD 2
PRINT HEAD 3
REGNTYPE = Q
Y
N
PERFORM QUAD REGN
PERFORM LINA REGN
LOAD Y, CHI LINES
LOAD, PRINT A LINE
LOAD, PRINT B LINE
REGNTYPE = Q
N
Y
LOAD, PRINT C LINE
PRINT Y, CHI LINES
REGNTYPE = B
N
Y
Q → REGNTYPE
TYPESETR

APPENDIX B

Listing

```
*
***********************************************************************
*
***  C:REGUSE     REGRESSION USE PROGRAM          C.E.T.          ^2/^2/7(
*
*
*    THIS PROGRAM PERFORMS A REGRESSION ANALYSIS UPON NOT MORE THAN 5^(
*    DATA POINTS, USING THE C:READ, C:REGN AND C:WRITE ROUTINES.  IT
*    YIELDS A WEIGHTED LEAST SQUARES FIT TO EITHER THE LINEAR EQUATION
*    Y=A+BX, OR THE QUADRATIC EQUATION Y=A+BX+CX**2, OR BOTH.
*    THE PROGRAM USES CARD INPUT AND LINE-PRINTER OUTPUT.
*
*
*         INPUT:
*
*         (1) A CARD BEARING L, Q OR B IN THE FIRST COLUMN, TO INDICATE
*             LINEAR OR QUADRATIC REGRESSION OR BOTH.  THE REMAINING 79
*             COLUMNS MAY CONTAIN THE USER'S TITLE.
*
*         (2) NOT MORE THAN 5^^ CARDS, EACH CONTAINING A SET OF VALUES OF
*             X, Y AND D.
*
*         THE ABOVE MAY BE REPEATED ANY NUMBER OF TIMES.
*
*         (3) A CARD BEARING # IN THE FIRST COLUMN, TO INDICATE THE END.
*
*
*         OUTPUT:
*
*         A, THE Y-AXIS INTERCEPT
*         B, THE COEFFICIENT OF X
*         C, THE COEFFICIENT OF X**2
*         SIGMA SQUARED A, THE INTERNAL ERROR (VARIANCE) OF A
*         SIGMA SQUARED B, THE INTERNAL ERROR (VARIANCE) OF B
*         SIGMA SQUARED C, THE INTERNAL ERROR (VARIANCE) OF C
*         S SQUARED A, THE EXTERNAL ERROR OF A
*         S SQUARED B, THE EXTERNAL ERROR OF B
*         S SQUARED C, THE EXTERNAL ERROR OF C
*         S SQUARED Y, THE EXTERNAL ERROR OF Y
*         CHI SQUARED
*
*
*     NOTES:
*
*     (1) X IS THE VARIABLE WITHOUT ERROR, Y IS THE VARIABLE WITH ERROR,
*         AND D IS THE STANDARD DEVIATION OF Y.
*     (2) EACH Y VALUE IS WEIGHTED BY THE RECIPROCAL OF THE SQUARE OF ITS
*         D VALUE;  THUS A FULL ARRAY OF POSITIVE D VALUES IS REQUIRED;
*         THE D VALUES MAY BE ONLY RELATIVE TO EACH OTHER.
*     (3) THE INTERNAL ERRORS ARE CALCULATED PURELY BY PROPAGATION OF THE
*         GIVEN STANDARD DEVIATIONS;  IF THE STANDARD DEVIATIONS ARE ONLY
*         RELATIVE, THE INTERNAL ERRORS ARE MEANINGLESS.
*     (4) THE EXTERNAL ERRORS REFLECT THE CLOSENESS OF FIT OF THE DATA.
*     (5) CHI SQUARED INDICATES THE AGREEMENT BETWEEN THE EXTERNAL AND
*         INTERNAL ERRORS FOR (N-M) DEGREES OF FREEDOM, WHERE N IS THE
*         NUMBER OF DATA POINTS AND M IS THE NUMBER OF PARAMETERS FITTED.
```

```
         REF      DECLREAD,INTGREAD,LINAREGN,QUADREGN,LFEWRITE
REGNTYPE RES      1
READTABL GEN,4,4,4,20,32  0,0,2,BA(READBUFF),80
READBUFF RES      20
TYPEBUFF RES      1
         BOUND    8
INDARRAY RES      1000
DEPARRAY RES      1000
DEVARRAY RES      1000
         BOUND    8
REGNTABL RES      22
CNTERROR GEN,16,16  24,0
         TEXT     '  ABORTED: TOO MANY DATA'
FMTERROR GEN,16,16  23,0
         TEXT     '  ABORTED: INVALID CARD'
REGUSEND GEN,16,16  14,0
         TEXT     '  END C:REGUSE'
LINHEAD1 TEXT     'I                                                     '
         TEXT     '       LINEAR REGRESSION'
LNHD1TBL GEN,4,4,4,20,32  0,4,2,BA(LINHEAD1),75
LINHEAD2 TEXT     '                                                      '
         TEXT     '              Y=A+BX'
LNHD2TBL GEN,4,4,4,20,32  0,4,2,BA(LINHEAD2),68
USERTITL TEXT     '1                                                     '
         TEXT     '                                                      '
         TEXT     '                              '
TITLTABL GEN,4,4,4,20,32  0,4,2,BA(USERTITL),132
HEADTHRE TEXT     'C                           PARAMETERS                '
         TEXT     '       INTERNAL ERRORS                         EXTERNA'
         TEXT     'L ERRORS'
HEAD3TBL GEN,4,4,4,20,32  0,4,2,BA(HEADTHRE),112
PARALINE TEXT     'A               A =                                SIG'
         TEXT     'SQA =                               SSQA =          '
         TEXT     '                  '
PARATABL GEN,4,4,4,20,32  0,4,2,BA(PARALINE),120
PARBLINE TEXT     '                B =                                SIG'
         TEXT     'SQB =                               SSQB =          '
         TEXT     '                  '
PARBTABL GEN,4,4,4,20,32  0,4,2,BA(PARBLINE),120
PARCLINE TEXT     '                C =                                SIG'
         TEXT     'SQC =                               SSQC =          '
         TEXT     '                  '
PARCTABL GEN,4,4,4,20,32  0,4,2,BA(PARCLINE),120
CHISLINE TEXT     'A                                                   CH'
         TEXT     'ISQ =                          '
CHISTABL GEN,4,4,4,20,32  0,4,2,BA(CHISLINE),80
SSQYLINE TEXT     '                                                      '
         TEXT     '                                    SSQY =          '
         TEXT     '                  '
SSQYTABL GEN,4,4,4,20,32  0,4,2,BA(SSQYLINE),120
POINTCNT RES      1
QADHEAD1 TEXT     'I                                                     '
         TEXT     '     QUADRATIC REGRESSION'
QDHD1TBL GEN,4,4,4,20,32  0,4,2,BA(QADHEAD1),76
QADHEAD2 TEXT     '                                                      '
         TEXT     '            Y=A+BX+CX**2'
QDHD2TBL GEN,4,4,4,20,32  0,4,2,BA(QADHEAD2),72
REGRNUSE LI,12    0                    REGNTYPE
         STW,12   REGNTYPE
```

```
CARDLOOP CAL3,1     READTABL         FETCH A CARD
         BCS,15     $
         LI,1       READBUFF         FOR INTGREAD AND DECLREAD
         LI,2       0                FOR INTGREAD
         LI,3       0                FOR INTGREAD
         BAL,0      INTGREAD
         BCR,2      DATAREAD         IF VALID NUMBER SYMBOL
         LB,0       READBUFF         CARD COLUMN 1
         STB,0      TYPEBUFF         FOR NEXT CALCULATION
         LW,12      REGNTYPE
         CI,12      0
         BCS,3      REGNHEAD         IF NOT FIRST RUN
TYPESETR LB,12      TYPEBUFF         REGNTYPE
         STW,12     REGNTYPE
         CI,12      '#'              END SIGNAL
         BCS,3      PRNTITLE         FOR NEXT RUN
         CAL4,15    REGUSEND
         LW,0       REGUSEND
         BCS,1      $-1
         CAL4,0     $                EXIT
PRNTITLE LI,1       79
         LB,0       READBUFF,1
         STB,0      USERTITL+2,1     LOAD TITLE (NOT COL. 1)
         BDR,1      $-2
         CAL4,10    10               FETCH TIME AND DATE
         STW,0      USERTITL+27
         STW,1      USERTITL+28
         STD,2      USERTITL+31
         CAL3,1     TITLTABL         PRINT TITLE, TIME, DATE
         BCS,15     $
         LI,6       0
         STW,6      POINTCNT
         B          CARDLOOP
DATAREAD LI,2       0                FOR DECLREAD
         LI,3       79               FOR DECLREAD
         MTW,1      POINTCNT
         LW,0       POINTCNT
         CI,0       500
         BCR,2      DECLCALL
         CAL4,15    CNTERROR         IF TOO MANY DATA
         LW,0       CNTERROR
         BCS,1      $-1
         CAL4,0     $                ABORT
DECLCALL BAL,0      DECLREAD         CONVERT X
         BCS,7      NUMERROR         IF INVALID NUMBER
         LW,6       POINTCNT
         STD,4      INDARRAY-2,6
         BAL,0      DECLREAD         CONVERT Y
         BCS,7      NUMERROR         IF INVALID NUMBER
         LW,6       POINTCNT
         STD,4      DEPARRAY-2,6
         BAL,0      DECLREAD         CONVERT D
         BCS,7      NUMERROR         IF INVALID NUMBER
         CI,4       0
         BCR,3      NUMERROR         IF 0
         LW,6       POINTCNT
         STD,4      DEVARRAY-2,6
         B          CARDLOOP
NUMERROR CAL4,15    FMTERROR         IF INVALID NUMBER
```

```
         LW,0      FMTERROR
         BCS,1     $-1
         CAL4,0    $                ABORT
REGNHEAD CI,12     'Q'
         BCR,3     QUADHEAD
         CAL3,1    LNHD1TBL
         BCS,15    $
         CAL3,1    LNHD2TBL
         BCS,15    $
         B         REGNSETR
QUADHEAD CAL3,1    QDHD1TBL
         BCS,15    $
         CAL3,1    QDHD2TBL
         BCS,15    $
REGNSETR CAL3,1    HEAD3TBL
         BCS,15    $
         LW,14     POINTCNT
         LI,15     INDARRAY
         STD,14    REGNTABL
         LI,14     DEPARRAY
         LI,15     DEVARRAY
         STD,14    REGNTABL+2
         LI,1      REGNTABL
         LW,12     REGNTYPE
         CI,12     'Q'
         BCR,3     QUADCALL
         BAL,0     LINAREGN         EXECUTE LINEAR REGRESSION
         B         PRINTOUT
QUADCALL BAL,0     QUADREGN         EXECUTE QUADRATIC REGRESSION
PRINTOUT LI,1      REGNTABL         TO CONVERT AND PRINT
         LI,2      BA(SSQYLINE)+97
         LI,3      BA(SSQYLINE)+118
         BAL,0     LFEWRITE
         LI,1      REGNTABL+2
         LI,2      BA(CHISLINE)+58
         LI,3      BA(CHISLINE)+79
         BAL,0     LFEWRITE
         LI,1      REGNTABL+4
         LI,2      BA(PARALINE)+17
         LI,3      BA(PARALINE)+38
         BAL,0     LFEWRITE
         LI,1      REGNTABL+6
         LI,2      BA(PARALINE)+58
         LI,3      BA(PARALINE)+79
         BAL,0     LFEWRITE
         LI,1      REGNTABL+8
         LI,2      BA(PARALINE)+97
         LI,3      BA(PARALINE)+118
         BAL,0     LFEWRITE
         CAL3,1    PARATABL
         BCS,15    $
         LI,1      REGNTABL+10
         LI,2      BA(PARBLINE)+17
         LI,3      BA(PARBLINE)+38
         BAL,0     LFEWRITE
         LI,1      REGNTABL+12
         LI,2      BA(PARBLINE)+58
         LI,3      BA(PARBLINE)+79
         BAL,0     LFEWRITE
```

```
         LI,1      REGNTABL+14
         LI,2      BA(PARBLINE)+97
         LI,3      BA(PARBLINE)+118
         BAL,0     LFEWRITE
         CAL3,1    PARBTABL
         BCS,15    $
         LW,12     REGNTYPE
         CI,12     'Q'
         BCS,3     SSQYPRNT
         LI,1      REGNTABL+16
         LI,2      BA(PARCLINE)+17
         LI,3      BA(PARCLINE)+38
         BAL,0     LFEWRITE
         LI,1      REGNTABL+18
         LI,2      BA(PARCLINE)+58
         LI,3      BA(PARCLINE)+79
         BAL,0     LFEWRITE
         LI,1      REGNTABL+20
         LI,2      BA(PARCLINE)+97
         LI,3      BA(PARCLINE)+118
         BAL,0     LFEWRITE
         CAL3,1    PARCTABL
         BCS,15    $
SSQYPRNT CAL3,1    SSQYTABL
         BCS,15    $
         CAL3,1    CHISTABL
         BCS,15    $
         LW,12     REGNTYPE
         CI,12     'B'
         BCS,3     TYPESETR
         LI,12     'Q'
         STW,12    REGNTYPE
         B         QUADHEAD
         END       REGRNUSE
```

ACKNOWLEDGMENTS

I am grateful to Paul Day and Arthur Jaffey for their helpful suggestions.

REFERENCES

1. Conrad E. Thalmayer, C:READ, A Reentrant Routine to Convert EBCDIC Decimal Numbers to Hexadecimal, ANL-7599 (Aug 1969).

2. Conrad E. Thalmayer, C:WRITE, A Reentrant Routine to Convert Hexadecimal Numbers to EBCDIC Decimal, ANL-7634 (Nov 1969).

3. Conrad E. Thalmayer, C:REGN, A Reentrant Linear and Quadratic Regression Routine, ANL-7660 (Jan 1970).

4. A. Jaffey, unpublished work.

x